BOG 6: ModSim Session 2

ASCR Workshop on Extreme Heterogeneity in HPC
23-25 Jan 2018

BOG 6 Contributors

Moderator(s): Jeremiah Wilke (SNL-CA), Zhiling Lan (Illinois Institute of Technology)

BOGists:

BOG 6: ModSim Themes from Whitepaper Submissions

- Performance Analysis Tools
  - On-node
  - Interconnect
- Post-Moore Scaling
  - 3D Stacking
  - ***Device models***
- Quantum Computing Models
- ML / Analytical modeling
  - Dark Silicon Management
  - Application Mapping
- Performance Models for Compilers
- FPGA Emulation Systems
  - Overlay
  - Memory / Accelerator Emulation
- Neuromorphic Processor models
- ***Smart network models***
- ***Scalable simulation frameworks***

Why are we here?

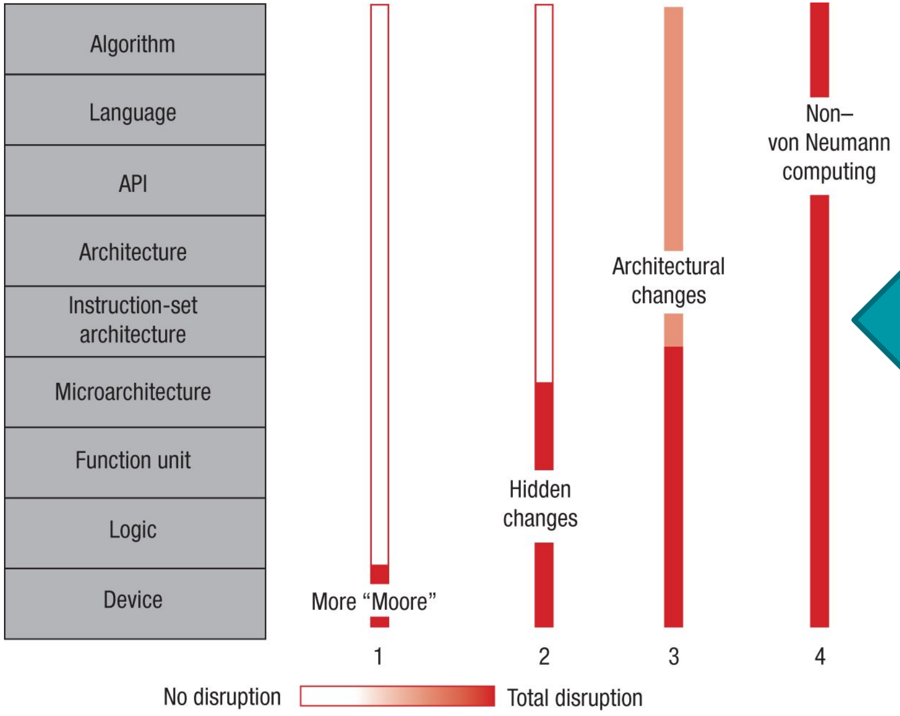
*Brainstorm and discuss what ModSim capabilities will be needed in the **2025-2035 timeframe** to make increasingly heterogeneous hardware technologies useful and productive for science applications - focusing on things that **industry will not solve** and problems that will be **post-ECP***

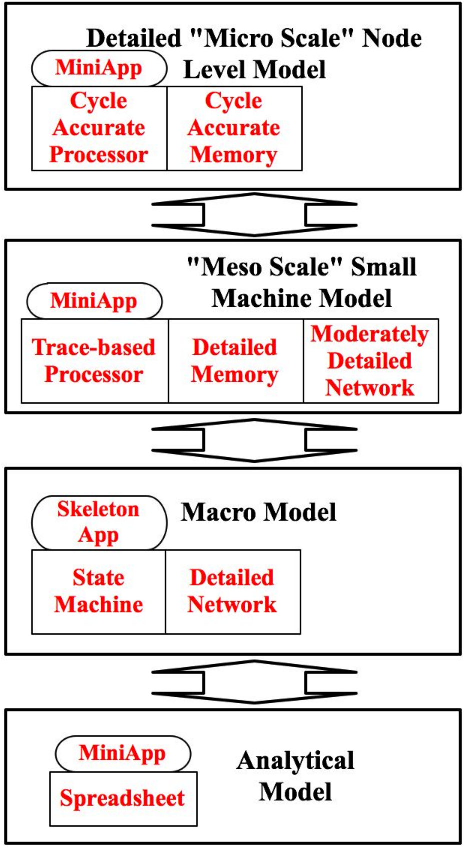
ModSim High-level Requirements:

- Quantitatively assess new architectural features
- Evaluate wide-range of granularities from microarchitecture to workflows
  - Emphasis on scalability
  - Understand interfaces/protocols in composing devices, system integration
- Characterize performance, power, resilience, and procurement cost tradeoffs
- Integrating components of novel accelerators (quantum, neuromorphic)
- Inform abstract machine models that provide cost models to:
  - Programming models and compilers
  - Algorithm development

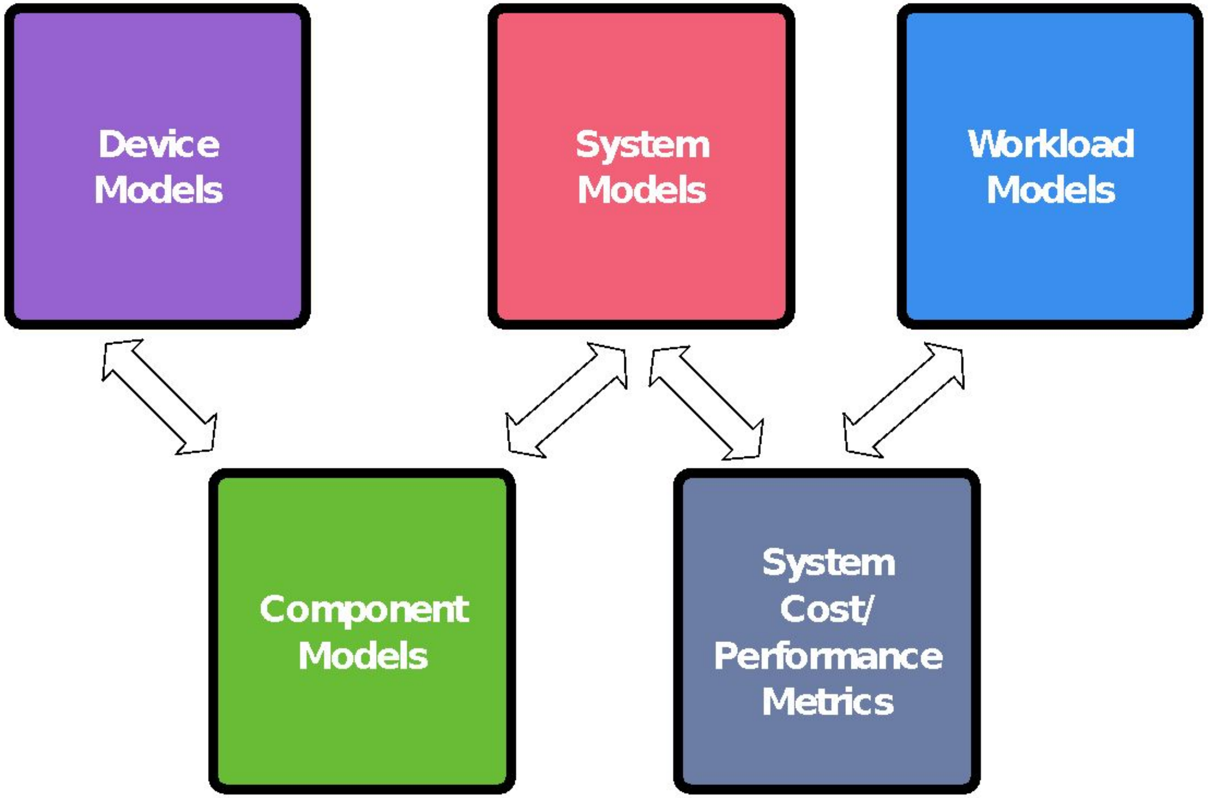
Goal for this breakout: **Identify N possible/promising research directions that address key challenges for DOE mission in the 2030+ timeframe. Focus on aspects related to heterogeneity.**

Mapping from disruptive technologies to ModSim research challenges





From micro to macro: Design parameters to system-level metrics



Review: Current state of system architectures and near-horizon architectures

- Homogeneous-heterogeneous:
  - Attached accelerators, but nodes overall uniform
  - Separate viz/analysis clusters
- Disruptive technologies on horizon for interconnects less extreme than on-node
  - Interconnect switches are already “custom accelerators”
  - Optics (photonics) being considered (particularly for big-data), but still part of overall electrical, packet networks
  - Smart networks, more logic in switches is research topic
- Increasingly hierarchical storage both for memory and storage
  - Burst buffers becoming standard, but exact design (e.g. placement of IO nodes) still an issue
- Workflow/architecture co-design
  - In situ/in-transit analysis driven by IO limitations
  - IO limitations driving new storage technologies

Review: Current state of simulation

- Simulation components organizing around common cores (SST, Codes, Manifold), but many components still stand-alone
- Many studies focused on ECP timescales, optimizing existing, soon-to-exist
  - Task placement and job scheduling of workloads
  - Routing, topologies (switches); protocols (NIC)
- Broadly speaking, two approaches dominate: analytical models and traces
  - Analytical modeling is fast, flexible, limited accuracy
  - Traces higher accuracy, but time-consuming and tied to existing platforms or extrapolation
    - High-fidelity modsim is time-consuming, sometimes resource-constrained
  - FPGA/Accelerator “behavior emulation” very efficient, scalability challenges (Florida)
  - ML “synthesis models” emerging in some domains, not widely used in HPC architecture ModSim
  - Lacking *dynamic, on-line* models
- Neuromorphic/quantum growing support both for single device, systems
  - Codes simulations up to 32K neurosynaptic cores
  - Integration of conventional and neuromorphic devices within SST
- Validation/verification/UQ: proposals, but no wide adoption
  - Need “trustworthy” models for UQ to be meaningful

Focus Questions: Bridge to Research Challenges

- What problems are industry going to solve? (out-of-scope)
- What problems are already part of ECP H&I? (out-of-scope)
- To what extent are system-level challenges an emergent property of node-level changes?
  - Accelerator/co-processors changing traffic patterns, network/IO provisioning requirements
  - E.g. NoC photonics changing cost/benefit of system-level photonics
  - E.g. On-node interconnects/protocols (e.g. GenZ) changing cost/benefit of disaggregated co-processors, memories
- How much focus is accelerating *current* workloads/apps? How much focus is accelerating *future/expected* workloads?
  - Separate paths for elephant/mice traffic major issue in big data systems, but not really discussed in many DOE capability systems
- How do we engage app/workflow/integration experts to make ModSim useful to them and keep ModSim from working in isolation?

ModSim Capability Targets and Research Directions

Challenge 1: Model integration

- Integration of heterogeneous device models into a system...
    - CPU, GPU, FPGA, memory, storage, network
- Integration of architecture models with heterogeneous software models...
    - e.g. batch scheduling models informed by hardware/software models
- Integration with heterogeneous cross-cutting service models managing power, resilience, thermal...

ModSim Capability Targets and Research Directions

Challenge 2: How to balance modeling speed, accuracy, and flexibility/extensibility for a “factorial” design space involving many application and device parameters?

- Analytical models…
- ML, synthesis models…
- Traces…
- Dynamic, online models…
- Behavioral emulation…
- Hybrid, multi-fidelity models...

Modsim session #2 Discussion 1/x

- coupling - need performance characterization I/O and workflow
  - Some are data driven, complexity is difficult to encapsulate in modsim
  - Reduce data from one type of accelerator, as an example of data driven; new challenges
- Dynamic model
  - Multi-scale modeling, split data streams, modsim of that would be useful
  - OS, DM sessions - predictability and performance portability - use modsim to build control models to use at higher level to manage at runtime; build control models out of lower level models
  - Dynamic models - has been a theme since modsim workshop 2012. Introspective, on the fly models. Become actionable via runtime model.
- Cost of simulation - tradeoff between accuracy and time to complete model

Discussion 2/x

- Node modsim concerns also extend to interconnect
- How to lower barrier to use modsim tools - to create model of workflow is too much work. Trace the workflow?
  - Codar ECP - run code and measure. Having a model to help resolve placement would be very helpful.
- Composability, scalability gets exacerbated at system level.
- Data sources can be greatly distributed - impact on facilities, applications,
- These challenges are similar to what are being faced now, not including wide area networks. Much commonality.
- Not only one way data streaming. Need quality of service concerns.
- Data movement, accelerators processing across network; now to access through software API is also a concern. Elevate level of abstraction of how to access - higher order objects. Have to support in modsim environment.

Discussion 3/x

- Need accurate model of complex workflows.
- In I/O have a common abstraction. Similar challenges.
- Simulations frameworks are simulating a stream of bytes - need to raise abstraction level in modsim.
- Accuracy of sim for quantum accelerator - how to integrate into system level simulator. There is no defined interface to integrate quantum accelerator model to system level. Model of interface - black box to abstract behavior?
  - Model cost of operations, same as today? Not necessarily design of quantum units affects quality of computation.
- Need synchronization of models for parallel modsim.
- Need interface to bridge between sim and analytic models.

Discussion 4/x

- How do take very expensive individual simulations and tie them into system level simulations
  - Need for physics models at system level.
  - "PCAD" models in electronic device models - these are too costly at that low level (3 - 4 devices, simulate materials). Want to cut down that level of sim, make it less accurate, but push to higher level abstraction. Device -> circuit.
  - Need to take simulators from accelerators and map up to system level
- What can we gain from industry big data centers models? What are differences between HPC centers and data centers.
- Validation is harder at system level - more moving parts. Higher variability.
  - High fidelity individual models but larger numbers of components in the mix.
  - Example - model memory controller to DRAM. but there was a gap when mem controller moved to CPU. Full system validation by BSC found 20% discrepancy of handoff from CPU core to mem controller. Just at one point in system.
  - Have to validate the entire system

Discussion 5/x

- Very accurate models of components, but feedback loops and dependencies are vitally important and must be modeled at system level.
    - Global address spaces - address dereference depends on conditions in global network.
- Variable precision of simulation
- Proprietary simulators that are more accurate are not available to incorporate into system model. Need synergy with vendors or will have to go with analytical models rather than cycle accurate.
- Functional correctness, performance matter a lot for accelerators. Security hole issues - bad actors can break in to an accelerator. Acc. makers need to supply functional and performance models so that we can find these vulnerabilities
- Will see emergent behavior in security and performance

Discussion 6/x

-

PRD 6.1 : Emergent behavior and validation

As we start combining individual simulation components, how do we ensure accuracy of the system as whole since components may be used in new (untested) contexts?

- Research challenges
  - Must define metrics for progress
- Potential research approaches and research directions
  - Must show how approach can be evaluated with progress metrics
- How and when will success impact technology?
  - Must answer why DOE needs to lead aside from industry

(Maybe several viewgraphs)

BOG 6:Key Research Challenges

Challenge 6.1: How do we create system-level workload models for extremely heterogeneous architectures that don't exist yet?
Challenge 6.2…

BOG 6: Possible Research Directions Summary

PRD 6.1 - direction 1

PRD 6.2 - direction 2

PRD 6.3 - direction 3

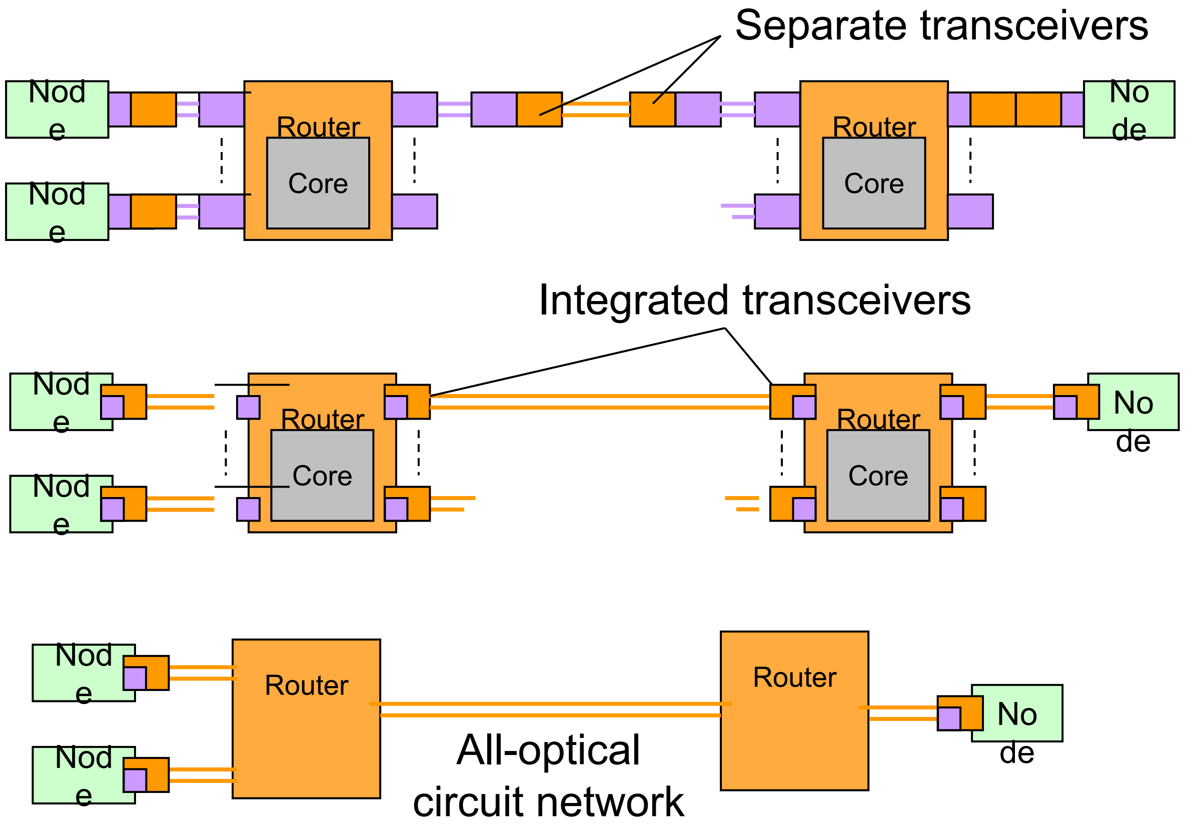
PRD 6.1 : Short title

One paragraph description (3 sentence/bullet)

- Research challenges
  - Must define metrics for progress
- Potential research approaches and research directions
  - Must show how approach can be evaluated with progress metrics
- How and when will success impact technology?
  - Must answer why DOE needs to lead aside from industry

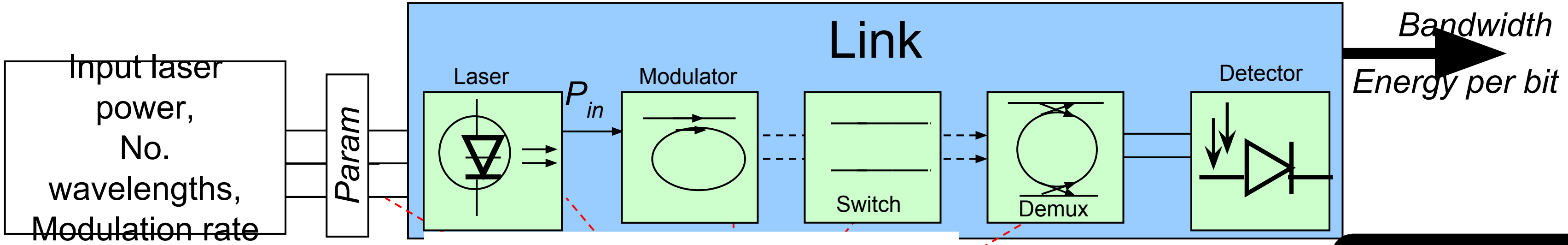
(Maybe several viewgraphs)

Silicon photonics example: Component model informed by abstract physical models

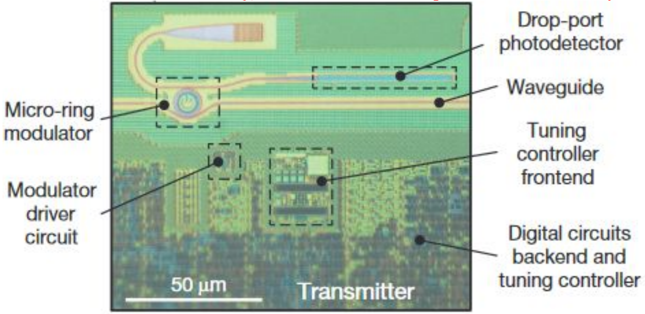


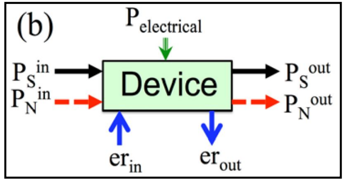
Component Models

Silicon photonics example: Abstract physical models must be developed from device models



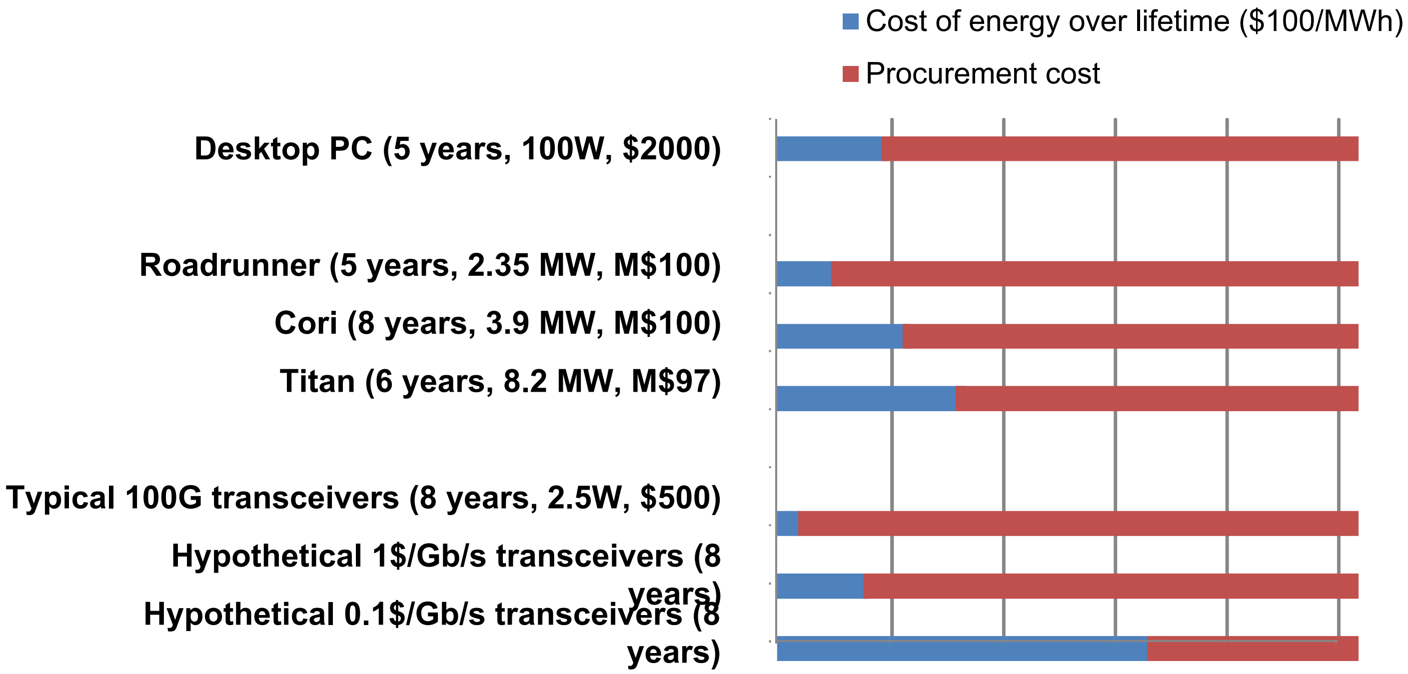
Abstract Physical Models





Device Models

Silicon photonics example: Component-level models will inform system-level cost/performance metrics



System Cost/ Performance Metrics